MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 3, 2022/2023

DSP5014 – STATISTICS AND PROBABILITY

(for diploma students only)

4 OCTOBER 2023

9.00 am – 11.00 am
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of 4 pages excluding the cover page and appendix.

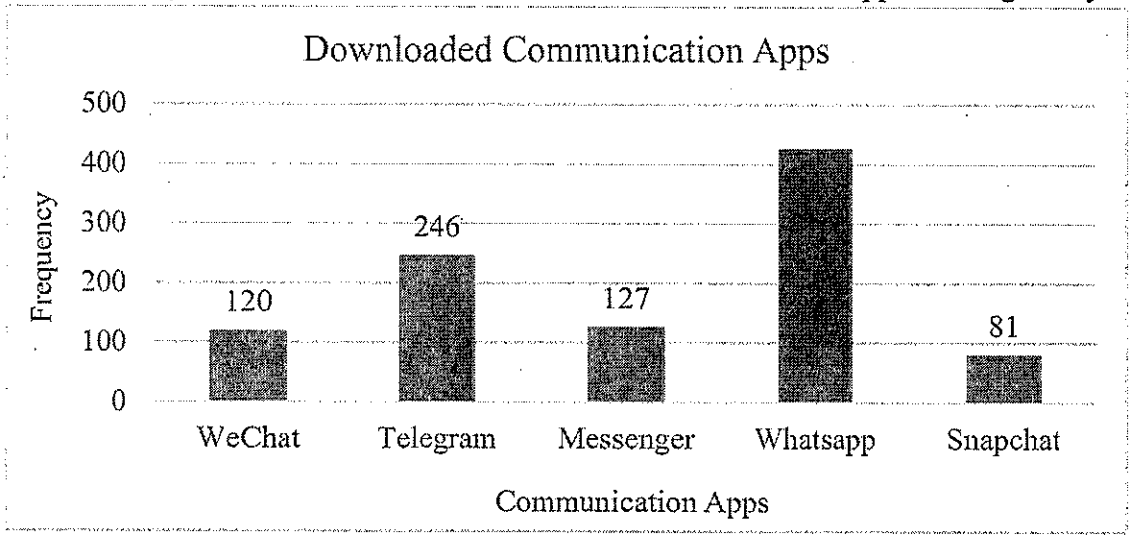
2. Attempt **ALL** questions. The distribution of the marks for each question is given.

3. Write all your answer in the Answer Booklet provided.

4. Key formulae are given in the Appendix.

Question 1

a) The following table shows the data on downloaded communication apps in Google Play Store.



i. Find the frequency of Whatsapp apps if the total number of downloaded communication apps is 1000. (1 mark)

ii. Represent the above information in a table which includes the category, frequency, relative frequency and percentage. (4 marks)

b) The following data shows a stem-and-leaf diagram that represents the time (in minutes) in MMU gymnasium over a period of one month among 50 diploma students.

Stems	Leaves
2	1 5 6 7 7 8
3	5 5 6 6 7 8 8 9 9
4	0 1 2 4 4 5 6 7 7 8 8 9
5	1 2 3 4 6 7 8 8 9 9
6	2 2 2 4 5 5 5
7	0 3 3 5
8	1 7

Key: 2 | 1 = 21

i. Determine the median and mode of the data. (4 marks)

ii. Construct a grouped-data-table for the stem-and-leaf diagram. The first class should start at $20 \leqslant 30$. (4 marks)

iii. From table in part (ii), calculate the sample mean and standard deviation. Round off the answer to one decimal place. (7 marks)

[TOTAL: 20 MARKS]

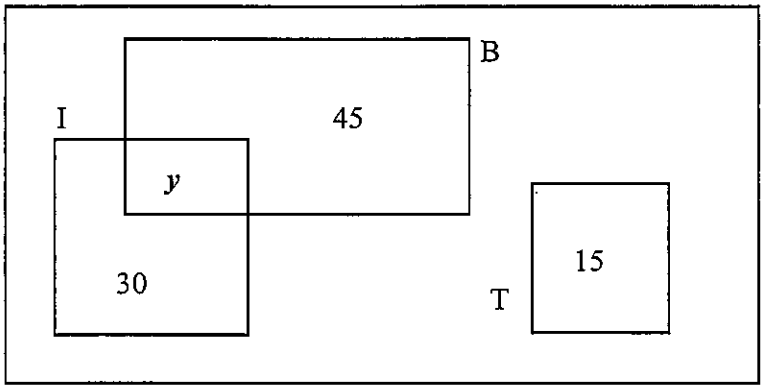
Continued...

Question 2

a) A factory has 3 assembly tracks that produce microchips. Track A produces 40% of the microchips and has a defective rate of 4%; Track B produces 30% of the microchips and has a defective rate of 3%; Track C produces 30% of the microchips and has a defective rate of 5%. A microchip is chosen at random from the factory.

i. Construct a tree diagram to represent the above information. (4 marks)
ii. Determine the percentage of defective microchips from Track A. (2 marks)
iii. What is the probability of getting the non-defective microchips? (5 marks)

b) The Venn diagram below shows a total of 100 kek batik cravers who bought the dessert for the past one month. The variety of kek batik are denoted as I – Indulgence, B – Bischoff and T – Tiramisu. One person is selected at random.



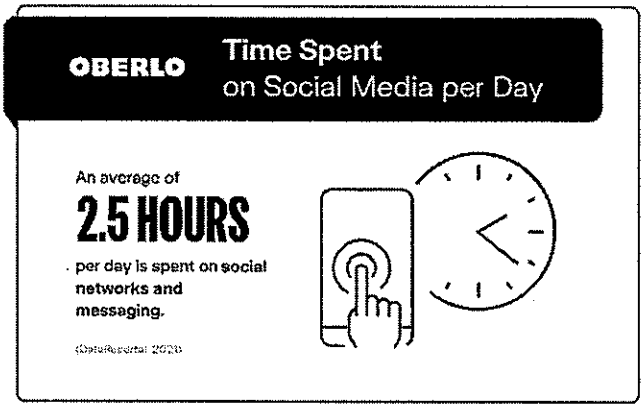
i. Identify the value of y. (1 mark)
ii. Find $P(I)$, $P(B)$ and $P(T)$. (3 marks)
iii. Find the percentage of a person who buys either Indulgence or Tiramisu dessert. (4 marks)
iv. Determine the probability of a person who buys a Bischoff dessert, given that he buys an Indulgence dessert. (4 marks)
v. Are the events *Bischoff* and *Tiramisu* mutually exclusive? State your reason. (2 marks)

[TOTAL: 25 MARKS]

Continued...

Question 3

a) From the DataReportal 2021, it shows that, an average of 2.5 hours per day is spent on social networks and messaging. A random sample of 100 users has been selected with a population standard deviation of 10 hours. Currently, the mean time spent per day on social networking is more than 2.7 hours. Assume that the data is approximately normal.



Source: https://www.oberlo.com/blog/social-media-marketing-statistics

i. Identify the *Ho* and *Ha*. (2 marks)
ii. Compute the value of test statistics. (3 marks)
iii. Determine the *p-value*. (2 marks)
iv. Test the claim at 10% significance level. (3 marks)

b) A survey concluded that, on average, customer spent less amount of money during the offline transaction compared to the online transaction. The data below summarize the results of the survey. Test the claim at 5% significance level. (10 marks)

Offline Transaction	Online Transaction
$\bar{x}_1$ = RM 2500	$\bar{x}_2$ = RM 2600
σ_1 = RM 500	σ_2 = RM 500
$n_1 = 250$	$n_2 = 250$

c) During a webinar session conducted via Zoom apps, 15% of participants faced an internet connection problem. It is shown that 40 out of 200 participants were unable to complete the webinar session due to an internet connection problem. Test the claim at $\alpha = 0.10$. (10 marks)

[TOTAL: 30 MARKS]

Continued...

Question 4

A practicum teacher is preparing a report to see the relationship between the number of missed classes, x and test scores, y for ten students in one private school. He presented the data in the following table.

Number of missed classes	1	6	5	2	4	3	1	5	2	1
Test score	80	f	55	60	45	55	78	45	67	85

a) Find the value of f if Σy is 600. (2 marks)
b) According to the table above, find the value of S_{xx} , S_{yy} and S_{xy}. (9 marks)
c) Compute the correlation coefficient, r and interpret the answer. (4 marks)
d) Find the linear regression equation $\hat{y} = \beta_0 + \beta_1 x$. (6 marks)
e) Compute the value of *SST*, *SSR* and *SSE*. (4 marks)

[TOTAL: 25 MARKS]

End of Paper.

APPENDIX – KEY FORMULA

	Ungrouped Data	Grouped Data
Mean	$\bar{x} = \frac{\sum x_i}{n}$ where n : sample size	$\bar{x} = \frac{\sum m_i f_i}{\sum f_i}$ where m : class midpoint f : class frequency
Variance	$s^2 = \frac{1}{n-1}\left[\sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n}\right]$ where n : sample size	$s^2 = \frac{1}{\left(\sum f_i\right)-1}\left[\sum m_i^{\,2} f_i - \frac{\left(\sum m_i f_i\right)^2}{\sum f_i}\right]$ where m : class midpoint f : class frequency

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \frac{x-\mu}{\sigma}$

HYPOTHESIS TESTING

Population	Mean	Proportion
1	$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$
2	$Z = \frac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}}$	$Z = \frac{\left(\hat{p}_1 - \hat{p}_2\right) - \left(p_1 - p_2\right)}{\sqrt{\hat{p}_p\left(1-\hat{p}_p\right)}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ *where* $\hat{p}_p = \frac{x_1 + x_2}{n_1 + n_2}$

LINEAR REGRESSION ANALYSIS

- **Least square regression equation,** $\hat{y} = \beta_0 + \beta_1 x$

  where $\beta_1 = \frac{S_{xy}}{S_{xx}}$ and $\beta_0 = \frac{1}{n}[\sum y - (\beta_1 . \sum x)]$

- **Correlation coefficient,** $r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$ where

$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$	$S_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$	$S_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$

- **Sum of Square**

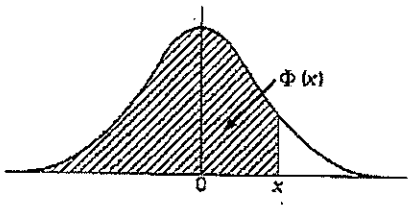
$SST = SSR + SSE$	$SST = S_{yy}$	$SSR = \frac{(S_{xy})^2}{S_{xx}}$

- **Test of Significance for Regression Slope**

Standard Error,	Regression t-Test	Correlation t-Test
$s_e = \sqrt{\frac{SSE}{n-2}}$	$t = \frac{\beta_1}{s_e / \sqrt{Sxx}}$	$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}$

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.



x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

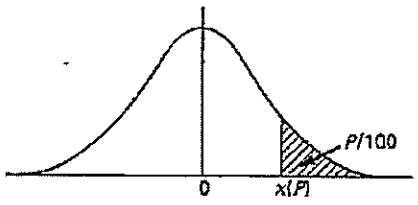
When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.



P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172